

# Chapter V: Style and multimodality

## V.1 Multimodality: a new approach to communication

In recent years, the topic of multimodality has gained growing attention in linguistics, education research, and a number of other disciplines (an overview is given in Jewitt 2014). Multimodal research, as well as the related tradition of multimedia research, allows for a look at different semiotic resources such as language, images, and music in their combinations and interactions. Multimodal communication poses new challenges for stylistics.

The term ‘mode’ is used for a ‘socially shaped and culturally given resource for making meaning’ (Kress 2014: 60). The concept is therefore related to that of ‘code’ (Eco 1986: 164–188) or ‘sign system’, which designates a system of signs that connects signifiers (expressions) with signifieds (meanings). However, there are a number of differences between these concepts, which cannot be explained in detail here, especially since all three terms are not defined and used consistently.

The matter is complicated further by a number of misunderstandings and simplified accounts, especially regarding the theoretical underpinnings of ‘code’ and ‘sign system’ in the semiotic tradition of Saussure and in structuralism. Suffice it to say that those speaking of ‘modes’ usually stress the role of contextual inference in interpretation and regard modes as resources without fixed meanings, while those speaking of ‘codes’ focus more on conventional aspects of meaning, and on the combinatorial (syntactic) restrictions imposed on signs organized in sign systems such as gesture or typography.

However, ‘mode’ is often used in a way that (implicitly or explicitly) includes verbal languages, sign languages, and signs in public places such as traffic signs – all of which are systems for meaning production that are largely conventionalized. It should also be noted that ‘codes’ or ‘sign systems’ are usually assumed to have a pragmatic dimension – the distinction between syntax, semantics, and pragmatics as three complementary aspects of sign systems was introduced by Charles Morris (1938: 6), see also Posner (1987: 25) – which accounts for some degree of context dependency of utterance meanings on the basis of conventional word meanings. The question of conventionality is therefore not sufficient to draw a clear-cut distinction between the terms ‘mode’ and ‘code’.

It is plausible, however, to apply ‘mode’ in a wider sense that includes the communication situation. This allows us to assume that codes are an optional aspect that occurs in some modes, namely those that encompass sign systems

with more or less conventional meanings, such as language or emblematic gesture. John Bateman has developed such an approach that proposes a definition of mode and relates it to an understanding of code as part of certain modes (Bateman 2011: 21). This description brings both terms into a systematic relation to each other, and thereby the different schools of semiotics that employ them.

From another perspective, Ellen Fricke has distinguished between various uses of the term 'mode' and the derived term 'multimodality'. A semiotic explication leads to a distinction of the two most frequent uses (Fricke 2012: 47). In some disciplines (e.g. psychology), 'mode' is primarily used in the sense of *perceptual mode*, and multimodality describes sign processes where various perceptual modes such as visual, auditory, and haptic perception are involved (Calvert et al. 2004). In other disciplines such as linguistics and film studies, 'mode' is more often used in the sense of *semiotic mode*. Multimodality then describes sign processes where various *sign systems* are involved. This leads to a wider understanding of multimodality because it comprises a number of additional cases, such as music and speech in a radio feature, where both modes are perceived auditorily, or images and written text in a comic strip as well as pictures and text in a book, where both modes are perceived visually.

Multimodality is a rapidly growing field in educational research, linguistics, and other areas, and empirical research on interactions between modes is an integral part of this field. However, research on the interplay of perceptual modes also has a long-standing history in cognitive psychology where it is often investigated with experimental methods.

Both traditions come together in experimental aesthetics, an interdisciplinary research field dating back to the beginnings of modern psychology in the 19<sup>th</sup> century that has recently seen a revival (as witnessed, for example, by the founding of the Max Planck Institute for Empirical Aesthetics in Frankfurt am Main in 2013).

In recent years, approaches to multimodal communication (e.g. Kress/van Leeuwen 2001; O'Halloran/Smith 2011; Bateman/Wildfeuer 2014) have received growing attention, sometimes with a specific focus on discourse analysis (Machin/Mayr 2012; Siefkes/Schöps 2013<sup>15</sup>). Multimodality research has begun to investigate the interactions between text and images (cf. Hess-Lüttich/Wenz

---

15 This special issue on new methods in discourse analysis includes, among its 10 contributions, 4 articles that present innovative approaches to multimodal discourse analysis, focusing on different modes and media.

2006; Bateman 2014), language and gesture (cf. Fricke 2013), and speech and moving images in film (cf. Bateman/Schmidt 2012; Wildfeuer 2013). This research lays the groundwork for a wide range of approaches to multimodality, and opens up new research routes.

It should be noted that multimodality research has a broad range of practical applications. Among them, the use of multimodal approaches for evaluating educational material and classroom situations (from primary school up to university and professional education) are one important area. The respective literature concerns itself both with the general consequences of multimodality for educational science (Saint-Georges 2013) and with practical perspectives for teaching (Unsworth 2011). Generally, the results of multimodality research are relevant to many professional contexts (for an overview cf. Jewitt 2014).

Examples for multimodal text types are combinations of image, graphics, and text (e.g. in newspapers, brochures, or websites), combinations of speech, gesture, facial expression, and body posture in face-to-face interactions, and combinations of moving images, speech, and background sound in video and film. As these examples show, multimodal texts are very widespread. In fact, depending on the definition of 'mode' and the granularity of distinctions, it can be argued that all communication is multimodal. Even a book that doesn't use images or graphics necessarily employs typographic means such as a specific page layout and fonts, or other semiotic resources such as page numbers, and structures the text in a specific way (through line spacing, page breaks, etc.), all of which influences our reading experience. Arguably, monomodal communication is rare or does not exist at all; certainly multimodal communication is the more common case.

These reflections have fundamental consequences for various disciplines, not least among them linguistics. Even today, analysing language on its own is probably the default (though no longer the only accepted) approach in linguistics. Multimodality research has shown that this is at best a limited perspective that cannot shed light on a number of important aspects of texts and communication. Worse, it could be claimed that monomodal approaches are methodologically invalid. If we are to believe the more radical assumptions in multimodality research, traditional approaches miss important aspects of every communication situation they analyse, and can therefore not result in adequate models of how language functions in the real world. Furthermore, the problem isn't solved by simply 'adding on' other modes in pragmatics (= the investigation of language use), where the context (and co-text) of language use is traditionally considered, while retaining the traditional models in syntax

and grammar. As Ellen Fricke has argued in detail, language integrates gesture even on the level of grammar, and a ‘multimodal grammar’ is needed in order to adequately describe these processes of syntactic and functional integration (Fricke 2012, 2013).

In a similar fashion, to focus on images, video, graphics, colour, or other modes regarded in separation from other signs surrounding them may only yield insights into what is – at best – a limiting case of human meaning-making. In various regards, linguistics has been woefully unprepared to deal with the complexity that results when the language-only (linguo-centric) view is given up. Linguistics still grapples to come to terms with the theoretical and methodological consequences of the multimodal revolution!

## **V.2 Multimodal stylistics**

At first glance, stylistics and multimodality are largely separate research traditions. Books on style usually look at one specific type of text or artefact, such as written or spoken language, images, fashion, architecture, design, or music. Similarly, in multimodality research ‘style’ is not always explicitly considered as a relevant category. This can probably be attributed to different research traditions and their respective terminologies, rather than to the assumption that stylistic phenomena are irrelevant for multimodal texts. In fact, a closer look reveals that a number of multimodal approaches do indeed use the term, whereas others investigate style-related phenomena without using terminology from stylistics.

It is important, in this context, that ‘style’ is a widely used term that is applied to quite different types of texts and artefacts, both in everyday language and in academic research (cf. Siefkes 2011, 2012a). It is this wide applicability of ‘style’ that recommends it as a suitable term for describing a common property of various modes in multimodal texts or artefacts. In fact, ‘style’ could be regarded as a category that is already multimodal, though this multimodality usually remains implicit. For example, while the style of images (both still and moving), language (both written and spoken), and music is often analysed, the relation of styles across these three modes of a film or video is rarely investigated.

It should be noted that multimodality research, in all its variation and different schools of thought, presupposes an underpinning that allows it to integrate the various modes that come together (e.g. music, images, and speech in a film), which are traditionally investigated in different disciplines. Multimodality research is therefore, by the very definition of its subject matter, an interdisciplinary undertaking. Stylistics, if we understand it as a general approach to

variation and individual expression in various areas, including music, images, language, architecture, design, and many others, can help to bridge the gap between these disciplines.

In a recent book on visual style, Stefan Meier explicitly points towards the importance of a multimodal approach to stylistics:

It has already been pointed out that the investigation of style in particular has to keep track of the multimodal correspondences between semiotic modes, in order to achieve a comprehensive understanding of meaning-making processes in communication. (Meier 2014: 188; translated from the German by MS)<sup>16</sup>

Meier (2014) proposes a model of style that is applicable in different areas, but his primary aim is to develop an adequate description of visual style. Meier's approach gives a central role to the selection of the content, which he regards as one aspect of style (Meier 2014: 201–211); further aspects are form-giving or formation (*Formung*; *ibid.*: 212–223) and composition (*ibid.*: 223–245) of the multimodal text. Meier understands his model of style as an explicitly multimodal approach to stylistic analysis (see also Meier 2012).

If we approach communication and texts as multimodal, this leads to new challenges as well as possibilities for stylistics. For example, if we analyse a website as a combination of images, written text, graphical elements (e.g. boxes, lines, and typographical decisions such as different font sizes), and navigational elements such as menus and links, we can analyse the style of these various modes separately.

---

16 [German original:] Es ist bereits verdeutlicht worden, dass gerade die Beschäftigung mit Stil immer die multimodalen Korrespondenzen mit anderen Zeichenmodalitäten im Blick behalten muss, wenn eine umfassende Behandlung kommunikativer Bedeutungstiftung vorgenommen werden soll.

Figure V.1: A website is a paradigmatic example of a multimodal text (from Pixabay, User JuralMin, Creative Commons CC0)



This gives us the possibility to compare them with other occurrences of the same mode across multimodal texts. For example, we can compare the use of typography, of images, or of graphics in printed newspapers and in online news websites. Multimodality research supplies us with the means to conduct a fine-grained analysis of style in certain contexts (such as websites or films) where otherwise only a generalized description would be possible. At least in principle, it allows us to uncover the contributions of the various modes to the overall stylistic qualities.

We have seen that there are different ways of considering the style of multimodal texts. An analysis might aim at a general stylistic description, or might start with considering style in the participating modes separately, and then look at the intermodal relations from a stylistic perspective.

Finally, it should be taken into account that various concepts of style that have been developed for linguistic style can, in some cases, usefully be expanded and applied to multimodal artefacts. One such concept is ‘mind style’, an idea that was introduced by Fowler (1977: 103). It was coined in order to describe stylistic

aspects of language used to highlight the mental life of a literary character. The concept has been quite successful in literary studies. It was taken up by scholars interested in cognitive approaches towards literature (e.g. Semino 2002), and has also been connected with corpus-based approaches (Semino/Short 2004) and computational methods (McIntyre/Archer 2010).

The analysis of 'mind style' is especially useful for understanding passages in literary works where the writing style seems to mimic the experiences made by a specific character, modulated by his or her mental capacities. Fowler analyses examples such as the mind style of the naive young woman Jenny Bunn in Kingsley Amis's *Take A Girl Like You* (1960) and of the Neanderthal boy Lok in William Golding's *The Inheritors* (1955) (Fowler 1977: 101, 105).

Rocío Montoro proposes an extension of the concept to other modes, resulting in a multimodal theory of mind style (Montoro 2010). Using the example of Bret Easton Ellis' novel *American Psycho* (1991) and the eponymous film (2000, dir. Mary Harron), he demonstrates that both the book and the film use specific (linguistic or visual) tropes to indicate the mindset of the main character, the yuppie and serial killer Patrick Bateman. In fact, the novel manages to convey that Patrick Bateman is a psychopath simply by using a specific style of writing in order to represent his thoughts. The style clearly conveys that he experiences his most brutal actions (murder and torture) in the same way as ordinary actions and events (Siefkes 2012a: 297).

A number of researchers have taken on the investigation of multimodal metaphor, which could be considered as a specific type of multimodal meaning (Stöckl 2004, 2010; Forceville 2003; Forceville/Urios-Aparisi 2009). An influential sub-field of multimodal metaphor theory, which has led to important advances in our understanding of multimodal meaning construction, is to be found in gesture studies. In human gestures (both conventional gesture and co-speech gesture), multimodal metaphor and metonymy play an important role in the integration of gestural and linguistic meaning (Parrill/Sweetser 2004; Kendon 2004; McNeill 2005; Fricke 2007; Müller 2008; Cienki/Müller 2008; Mittelberg 2010).

Generally speaking, multimodal metaphor is rarely brought into connection with the concept of style, possibly because cognitive semantics – the linguistic tradition which has exerted a strong influence on current multimodal metaphor research – seems reluctant to include the term 'style' in their theories. Even though many researchers don't explicitly formulate their findings in the terminology of stylistics, they often focus on aspects that have stylistic relevance. Traditionally, metaphor and metonymy have been regarded as stylistic devices, and their use certainly influences the overall stylistic impression of a multimodal artefact or text. Multimodal metaphor, as well as multimodal metonymy

(cf. Forceville 2009; Mittelberg/Waugh 2009), should be considered as a part of multimodal meaning that is closely connected to questions of style in multimodal texts. Neither traditional stylistics nor conventional metaphor theory suffice to explain the phenomena of multimodal metaphor and metonymy, because they do not allow for the interaction of different modes in the construction of metaphorical meaning. Since traditional metaphor theories are often rooted in linguistics and difficult to apply to non-linguistic domains, multimodal theories of metaphor are usually developed on the basis of cognitive linguistics and conceptual metaphor theory (cf. Stöckl 2010).

Fricke (2006) proposes an approach towards multimodal metaphor based on Fauconnier's and Turner's Mental Space Theory, also called (Conceptual) Blending Theory in its more recent incarnations (Turner/Fauconnier 2000; Fauconnier/Turner 2003). Fricke explains how mental spaces have been used to construct metaphorical meaning in Georg Nussbaumer's *orpheusarchipel*, an opera in the form of an installation. *Orpheusarchipel* used a three-story bunker to represent the elements of the Orpheus myth metaphorically in relation to the human body and its cognitive, emotional, and physiological functions. The specific use of intermedial (and in some cases also intermodal) integration on the basis of mental spaces is characteristic for Georg Nussbaumer's multimodal style, as Fricke (2006: 138) points out.

In their book *English in Urban Classrooms*, Kress et al. (2005) investigate, among other topics, how teaching style influences the classroom and the learning success of students, modulated by the different methods and approaches used. They studied the teaching methods of individual teachers in English classrooms and found marked differences in teaching styles. They point out that these are connected to classroom layout, and the use of classroom space by the teacher and the students during class. They understand the English classroom as a multimodal sign, and analyse different examples of how it is connected to styles of teaching (Kress 2005: 21–36), and how it shapes the learning style of students. The classroom can be arranged and used in different ways that either foster or block certain kinds of interactions. Different classroom layouts are connected with (implicit or explicit) educational principles, and ideas about how the classroom should function as a space for communication and learning.

An individual teaching style, in the sense of everything that is typical for one teacher and that distinguishes the experience of being taught by her or him, can be analysed into a number of more specific styles, such as the arrangement of space in the classroom and the style of using it during teaching, teaching style in the more specific sense of pedagogy, the stylistic qualities of the teacher's speech,

and interaction styles of teacher and students. One relevant aspect is the use of different modal resources that are employed during teaching, including spontaneous spoken language, writing on the blackboard, novels, journals and other illustrated printed documents, group work and discussions, films, computer and tablet use, etc. A number of examples analysed by Kress et al. (2005: 37–68) show that the selection from these multimodal resources, the intensity of their use, and their combination in various learning situations in the English classroom differ significantly between teachers. The use of modes can therefore be regarded as one aspect of teaching style, and there can be little doubt that they influence the learning styles of students.

### **V.3 Intermodality**

If a text or discourse uses various modes, these will obviously not be regarded independently of each other. Rather, the text combines them and forms them into a coherent textual whole. For example, in face-to-face communication the modes speech, gesture, facial expression, and body posture form an integrated stream of communicated meaning which cannot be explained simply as an addition of the meaning conveyed in the separate modes. Modes can enter in various kinds of relations to each other, for example contrast or complement each other, can be used to emphasize something that is conveyed in another mode, etc.

Some differences in terminology exist and should be noted. Whereas in linguistics and semiotics, the terms ‘intermodality’ and ‘intersemiosis’ are frequently used (e.g. Wildfeuer 2012), psychological studies often speak of ‘cross-modal interactions’ (e.g. Vines et al. 2011).

A number of systematic models and schemata for describing and annotating relations between modes have been developed (Royce 1998; Oviatt 1999; Wildfeuer 2012; Siefkes 2015). Many approaches focus on specific areas of intermodality, such as image-text relations (Marsh/White 2003; Martinec/Salway 2005; Liu/O’Halloran 2009; Bateman 2014) or on the influence of music on the perception of film (Pavlović/Marković 2011; Cohen 2013). Siefkes (2015) proposes a fully general model of intermodal relations.

#### **V.3.1 Style in different modes**

How does style function in a multimodal text or artefact? It is plausible to assume that style can be described separately for at least some (and maybe for all) modes that make up a multimodal text. However, the stylistic description of the

separate modes will not yield an adequate and complete description of the style of a multimodal text or artefact.

Let's take an example. If we think about the style of a feature film from the perspective of multimodality research, it becomes clear that the primary modes can be investigated separately regarding their stylistic qualities. In many cases, the moving images, the characters' speech, and the music will each have their own, distinctive style. However, it is equally obvious that the results of the separate analyses cannot be the last word on the film's style. Rather, we have to consider how an overall stylistic impression is created on the basis of the separate styles, and which holistic style effects may exist that cannot be reduced to contributions of the separate modes.

The relations between style and content in the various semiotic modes of a multimodal artefact can influence the perception of the whole. For example, in a film or music video, the visual, verbal, and musical style can influence the perception of the other modes, and contribute to a complex and interesting overall experience. This is especially relevant for aesthetic texts or artefacts, where differences between the modes can lead to additional complexity and give rise to further thoughts and interpretations – as long as it is possible to bring the modes into relation to each other, and interpret them as contributing to an artistic whole.

Unfortunately, relatively little work has been done on the stylistic dimension of intermodal relations. While many approaches to intermodality or intermediality (cf. Elleström 2010) do not expressly limit their perspective to questions of form and content, style is rarely considered as one aspect of modes that can also interact with other modes, although it is not hard to find examples for such intermodal stylistic influences (e.g. in websites, films, or illustrated books). Siefkes (2015) therefore proposes to consider style as a separate stratum that can interact with the style (and potentially with other dimensions, especially the content, and possibly also the form) of other modes. This approach, which explicitly includes the stylistic dimension in a formalized model of intermodal relations, is one of the first detailed theories of stylistic intermodality.

Of course, the relevance of intermodality, and the specific effects that can be attributed to intermodal relations, depends on the type of multimodal texts or artefacts that are investigated. Intermodal relations have different effects for various types of multimodal artefacts such as websites, illustrated textbooks, or films. It is therefore necessary to look at specific cases in order to understand how modes influence each other, in regard to both content and style (Siefkes 2015: 118).

The investigation of style in multimodal texts raises the question of styles in different modes, and their interaction. Different approaches to this problem can

be taken. It is possible to regard multimodal texts or artefacts from a 'holistic' perspective and investigate only the style of the textual whole, integrated across all participating modes. In this holistic view, a multimodal text can be compared with other texts that employ the same modes (e.g. a film with other films, a website with embedded videos with similar websites, etc.), but it doesn't make any sense to ask for the style of the separate modes. However, this perspective is unnecessarily narrow because it precludes the possibility of looking at the modes separately. It also makes it very difficult to compare multimodal artefacts which do not make use of the same modes.

These limits can be avoided by including a stylistic analysis of the separate modes, which enables the researcher to look at contributions of the modes to the overall stylistic qualities and effects. This allows researchers to compare multimodal texts that have some, but not all modes in common, by 'adding' or 'subtracting' contributions of additional or missing modes to the overall stylistic properties. A website with embedded videos can now be compared with one that doesn't include videos, a comic that uses speech bubbles with one that doesn't employ them, etc.

Such comparisons will not always lead to relevant results, but they are helpful in cases where it is plausible that the contribution of specific modes (which might be regarded as 'optional' in a certain multimodal text type or genre) can be factored out of the overall stylistic configuration. For example, films might be compared to each other in regard to their film style even when one of them does not use a soundtrack, by looking at the visual style as well as the style of speech and the use of background sound.

### **V.3.2 Stylistic relations between modes**

A separate investigation of the style of modes has to be supplemented by looking at the relations between the styles of different modes. In multimodality research, relations between modes are investigated in the subfield of *intermodality*.

For the special case of image-text relations, an approach towards intermodal relations that considers style as one aspect of intermodality has been developed by Emily Marsh and Marilyn White (Marsh/White 2003). They propose a taxonomy of image-text relations formulated as functions that create a specific relationship between the modalities. Some of these functions refer to stylistic qualities, most explicitly in the function *Match style*, which is present when 'image and text match along some stylistic dimension' (ibid.: 667). Some of the functions proposed by Marsh and White define image-text relations that can apply either to style or to other aspects of modes. For example, the function

*Elicit emotion* is present when the images ‘encourage emotional response from [the] reader through display of content or style that is especially arresting or disturbing’, and the function *Alienate* means to ‘create tension between image and text through contrast in style or mood’, but explicitly excludes content relations, for which the comparable intermodal relation *Contrast* is introduced (ibid.: 667).

Figure V.2: A taxonomy of image-text relations (from Marsch/White 2003: 653)

A Functions expressing little relation to the text	B Functions expressing close relation to the text	C Functions that go beyond the text
<i>A1 Decorate</i>	<i>B1 Reiterate</i>	<i>C1 Interpret</i>
A1.1 Change pace	B1.1 Concretize	C1.1 Emphasize
A1.2 Match style	B1.1.1 Sample	C1.2 Document
<i>A2 Elicit emotion</i>	B1.1.1.1 Author/Source	<i>C2 Develop</i>
A2.1 Alienate	B1.2 Humanize	C2.1 Compare
A2.2 Express poetically	B1.3 Common referent	C2.2 Contrast
<i>A3 Control</i>	B1.4 Describe	<i>C3 Transform</i>
A3.1 Engage	B1.5 Graph	C3.1 Alternate progress
A3.2 Motivate	B1.6 Exemplify	C3.2 Model
	B1.7 Translate	C3.2.1 Model cognitive process
	<i>B2 Organize</i>	C3.2.2 Model physical process
	B2.1 Isolate	C3.3 Inspire
	B2.2 Contain	
	B2.3 Locate	
	B2.4 Induce perspective	
	<i>B3 Relate</i>	
	B3.1 Compare	
	B3.2 Contrast	
	B3.3 Parallel	
	<i>B4 Condense</i>	
	B4.1 Concentrate	
	B4.2 Compact	
	<i>B5 Explain</i>	
	B5.1 Define	
	B5.2 Complement	

The taxonomy proposed in Marsh/White 2003 does not systematically include style, and only mentions it in regard to some intermodal relations.

In Siefkes (2015: 118), it is proposed that style be understood as one of three strata of all texts, the other two being expression (= form) and semantics (= content). On this basis, a number of intermodal relations can be used to describe how the stylistic properties of different modes interact with each other. Furthermore, the style of one mode can also interact with aspects of form or content of other modes, i.e. with non-stylistic aspects.

For example, the relation *Intermodal Emphasis* describes the case where aspects of one mode emphasize aspects of other modes. Siefkes (2015) explains this relation with an example from the film *Gattaca* (1997, dir. Andrew Niccol):

[This example] analyses a film scene from ‘Gattaca’ in which two protagonists have a swimming competition in the open sea at night. Anton, one of the competitors, yells ‘Where’s the shore? We’re too far out!’ Shortly afterwards, Anton, who is exhausted, is shown as he sinks below the surface and nearly drowns. The visual style of the scene employs shaky hand camera movements, presents the scene from angles both above and below the water, and changes between close-ups and distance views. The moon’s reflection on the waves provides erratic lighting.

In this example, the character’s speech already characterises the situation as dangerous. The visual style and the lighting style support and strengthen this interpretation. The *Intermodal Emphasis* IRT (cf. Marsh/White 2003: 653) can be inferred to hold between both the visual style  $v.i^{sty}$  and the style of lighting  $v.lighting^{sty}$  as emphasising modes, and auditory language  $a.l$  as the emphasised mode. (Siefkes 2015: 126)

This is just one example for how the style of one mode can be analysed in its relations and interactions with both stylistic and non-stylistic aspects of other modes. It is therefore important to explicitly include style in models of intermodality.

To summarize, intermodality is not limited to stylistic properties, but intermodal relations can also be found between styles (e.g. when a text is written in a clear and simple style, and accompanied by images that can be characterized as clear and simple). Intermodal stylistic relations should be regarded as one aspect of intermodality, just as intermodal relations exist on the level of expression (= form), e.g. when a text is overlaid over an image, or of content, e.g. when a text contradicts an image (Siefkes 2015: 118). Intermodality research can point towards connections between the styles of different semiotic modes (e.g. of a verbal text and the accompanying images), or between the style of one semiotic mode and aspects of expression or of content of other modes (e.g. between a visual style and the content of a verbal description).

### V.3.3 Experiments on intermodality

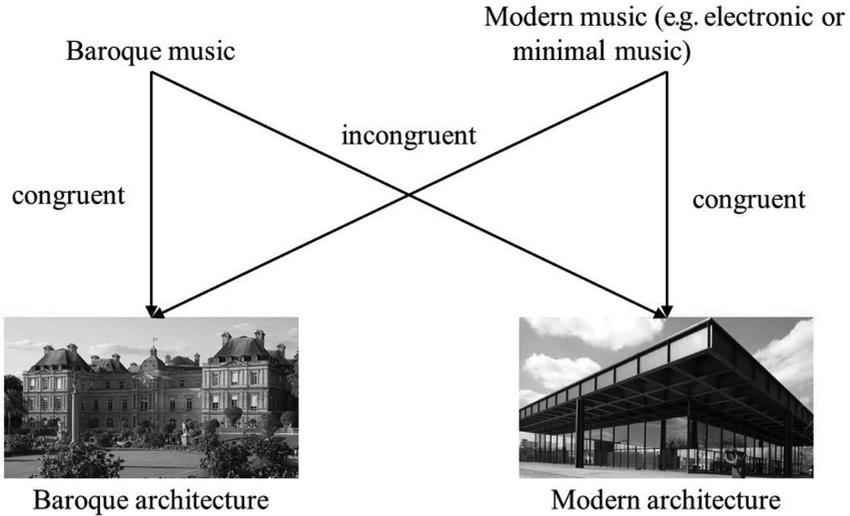
As we have seen in the last section, there is a substantial body of work on intermodal relations. The specific question of intermodal (or cross-modal) influences of style, however, has received comparatively little attention.

The authors of this book investigated intermodal stylistic relations in experimental settings. We conducted two studies at the University IUAV of Venice, which focused on the interactions between the aesthetic perception of architecture and background music in audio-visual presentations (Siefkes/Arielli 2015).

Specifically, it was investigated how two musical styles (modern and baroque) influence the judgment of buildings belonging to the corresponding architectural styles. The studies connected two different types of artefacts, namely buildings and pieces of music, both of which are connected with meanings and associations and can therefore be regarded as semiotic modes, and are also perceived in different perceptual modes, namely visual and auditory perception. It was examined how music influences the aesthetic judgment of buildings belonging to the corresponding architectural styles.

The results for the influence of background music were inconclusive. However, both studies demonstrated the effect of *intermodal congruence* or *incongruence* on the judgment of buildings. When the music was congruent with the architecture (e.g. baroque architecture and baroque music), the observed architectural style was judged as more ‘balanced’, more ‘coherent’, and in the second study also as more ‘complete’ (Siefkes/Arielli 2015: 261). Notably, the subjects were explicitly asked to consider the style of the architecture, not the music, and did not intentionally rate the relationship between music and architecture.

Figure V.3: Effects of intermodal congruence vs incongruence (Siefkes/Arielli 2015)



Generally, it should be noted that intermodality, the consideration of relations between modes, is an important and still under-researched topic. In practice,

multimodal analyses often concentrate on the separate semiotic modes, or assume that meaning is produced by all modes together, licensing researchers to draw on each mode at will when explaining the overall meaning. However, both of these alternatives are problematic.

The approaches detailed above, and the examples given for intermodal relations, show the relevance of intermodality. Apart from mode-specific contributions and holistically produced meanings, specific relations between semiotic modes play a significant role in the creation of multimodal meaning. Accordingly, they should be taken into account in multimodal analyses. It is therefore important to understand which types of intermodal relations can be assumed, and to integrate them into a general model of text/discourse analysis that adequately represents all participating modes and textual levels.

### **V.3.4 Intermodality in film and video**

One area of intermodality that has been comparatively well-researched is film and video. For example, the film theorist Edward Branigan gives an example of how camera style influences the viewer's attention to aspects of the film:

A camera movement that is motivated by the narrative is relatively 'invisible' as an element of style with the result that plot and story are foregrounded for a spectator. By contrast, an unmotivated camera movement draws attention to itself in relation to plot and story; or rather, it will seem to a spectator that there has been some kind of disturbance in the plot – a mysterious ellipsis, pause, expansion, digression, break, or obscurity. These effects are important even if a camera movement is later partially motivated or motivated retrospectively. (Branigan 2006: 26)

Under the term 'camera movement', Branigan subsumes the position, angle, and focus of the camera (Branigan 2006: 25). Camera movement defined in this way determines a large part of a film's visual style. If camera movement is congruent with story and plot, it becomes more or less invisible. A related concept can be found in the area of film editing, where the term 'continuity editing' is used to describe a style that is intended to make the editing and cutting of the film invisible, and let the viewer focus on the narrative (Bordwell/Thompson 2008: 231–251). Continuity editing is a camera technique that relies to some degree on intermodal congruity: for example, cuts have to be more or less in accordance with character's speech. If a close-up continues after a character stops talking, this would draw attention to the editing process, and would prompt the inference that there is something the viewer should pay attention to (e.g. the facial expression of the character).

Hartmut Stöckl (2003) proposes a multimodal approach to style analysis based on the model of Kress/van Leeuwen (2001). He uses the example of a television commercial to develop a fine-grained analysis that considers style as the specific uses made of modes and of their relations to each other. The level of the whole text, the most important modes (language, images, and music), and intermodal relations between these modes are considered. The basis for the analysis is a fine-grained transcription of the multimodal text (ibid.: 313). For the mode *language*, the approach distinguishes between thematic structure, rhetoric, lexis (vocabulary), prosody, and typography (ibid.: 315–316), for *images*, between image content (*Bildinhalte*), image design (*Bildgestaltung*), and narrative structure (ibid.: 316–317), and for *music*, between melody, instrumentation, volume, as well as the connotations connected with a specific musical piece and its composers (ibid.: 318).

Stöckl's approach integrates many important dimensions and adequately considers the complexity and internal structure of modes, as well as intermodal relations and textual structures. It is not always obvious how stylistic properties can be separated from other categories; in fact, the system seems to include some content aspects as well. Stöckl (2003: 311) points out that no features can be discounted *a priori* as irrelevant for multimodal style. However, a later article demonstrates a comparable multimodal analysis of a television commercial where the use of 'style' is more limited (Stöckl 2011). Generally, it should be considered how the large amount of analysed detail contributes to the more general perception and judgment of style. We should search for *principles of salience* that describe which (combinations of) aspects contribute to our overall perception of a style. Aspects may become salient when they are unusual, interact with other aspects by strengthening or contradicting them, or are relevant to the function or context of the multimodal text.

Generally, styles can rely to different degrees on the various modes used in a multimodal text, or even on specific intermodal relations between two or more of the modes:

[W]e characterized Robert Bresson as a director who makes sound particularly important in many of his films; we analyzed several important ways in which sound related to image in *A Man Escaped*. This use of sound is one aspect of Bresson's unique style. Similarly, we looked at *Our Hospitality* in terms of how its comic *mise-en-scène* is organized around a consistent use of long shots; this is part of Buster Keaton's style in other films, too. (Bordwell/Thompson 2008: 304)

We can therefore conclude that some aspects and techniques of film style, such as camera or editing styles, depend on intermodal relations between images, characters' speech, and music. However, film style also includes how the modes are

employed to tell the story, and combined to form a textual whole. These aspects of film style can be generalized towards all areas of multimodal style. A specific way of delegating parts of the narrative to different modes, and using intermodal relations to bind them together, can be understood as a ‘style of using modes’, to distinguish it from the more general concept of ‘multimodal style’ which also comprises mode-specific styles and holistic style effects.

Janina Wildfeuer (2012, 2013) has developed a linguistic approach towards film analysis that builds on Segmented Discourse Representation Theory (Asher/Lascarides 2003). Her analysis aims towards a dynamic account which understands a film as a ‘text’ or ‘discourse’ in which meaning is constructed by the viewers over time, on the basis of textual clues. This analysis focuses on the production of filmic coherence through the various semiotic modes and their interactions:

Meaning in film arises out of the multiple interaction of various modalities such as images, sounds, music, gestures, camera effects, etc. [...] The interplay of the modalities results in a narrative *text* whose comprehension and interpretation requires the spectator’s active participation. As a dynamically unfolding discourse, the combinations of resources change in time and space and continually produce meaningful sequences which have to be analysed during their progression. [...] Contemporary approaches to film analysis have not yet succeeded in systematically defining and describing how filmic devices are intersemiotically combined to narrative structures. This book investigates exactly how this interpretation process takes place with regard to the overall *coherence* of the text. (Wildfeuer 2013: 1; emphasis in the original)

In one chapter of his book *Analysing Popular Music*, David Machin shows how the style of music videos modulates the message given by the music and the lyrics. For example, the music video to The Clash’s *London Calling* from 1979 shows images from London in dark tones:

Beginning with a shot of Big Ben at the Houses of Parliament we see a number of scenes cast in dark shadows, or washed out with light [...]. There are fast edits with short sequences. [...] But as a composition its uses of extremes of tone of dark and shadow heavily stylise the film into a noir piece. [...] While the music of London calling is certainly lively, it is tempered by the dark moods and blinding unkind lights. (Machin 2010: 195)

In this example, it could be argued that the video not so much adds an additional dimension of meaning, but rather clarifies or ‘disambiguates’ (Siefkes 2015: 125) how we have to understand the music. While the uptempo beat, catchy melody, and overall danceable quality of this classic punk-rock song connote a lively mood and a certain willingness to entertain, the images of the official music video immediately produce an atmosphere of danger. During the intro of the

song, the stark contrast between light and dark, as well as the sometimes jerky images of the hand-held camera and the aggressive movements of the musicians while they walk, run, or play their guitars, contributes to the threatening atmosphere and the impression of hardly contained aggression. This general feeling is confirmed when the lyrics start with the lines 'London calling to the faraway towns/ Now war is declared – and battle come down'. Therefore, the images and spoken language (lyrics) support each other in this video, and jointly disambiguate the music of the song towards a dark and aggressive feeling anchored in its raw energy and uptempo beat.

While this analysis captures important aspects, the relations between different modes in music videos – as well as in other multimodal artefacts – are often more complex. While we may aptly characterize relations between two modes, for example, in terms of 'similarity' or 'contrast', usually the various modes contribute in more complex ways to the overall meaning. David Machin goes on to characterize how as a teenager, he experienced the music of The Clash as rebellious, but later became unsure if the band actually had a message:

But what The Clash stood for is not clear. Lyrically they hint at rebellion, although they don't say against what. Musically their music is tense and lively, with stern singing, sometimes confiding and at others yelling angrily. In videos, such as for 'London calling', we find them cool and slightly mysterious, inhabiting sleazy environments. So the video, sounds and words are littered with connotations of certain discourses. (Machin 2010: 195)

Machin hints at the important point that the relations between modes may be quite indirect, but still relevant for our perception. While it is tempting to look at relatively simple relations between modes, such as similarity or contrast, in fact the connections may often be more indirect and may happen through allusions to discourses, stereotypes, or other forms of shared knowledge. While the images of London and of the band performing near a river (probably the Thames) have *prima facie* nothing to do with the lyrics, which speak of 'war', 'meltdown', 'zombies of death', and 'nuclear error', they can be connected via discursive stereotypes such as the drabness and cruelty of the inner city, and the outsider who 'live[s] by the river' because he sees society as doomed and expects the end of Western civilization.

As this example shows, relations between different modes can be quite indirect and may, for example, be constructed through allusions to discourses or stereotypes (Siefkes in print). Even in this case, they contribute to the overall artistic impression of the work. In fact, it may even be advantageous for our feeling of artistic unity if the relations between modes are less direct, because this allows for more freedom of interpretation, giving us the opportunity to create our own unity out of the diverging meanings that we encounter.

The interpretation of an artwork such as a music video may be seen as more interesting if we can guess at complex connections between images, music, language etc., without being able to put our finger on which connections these are, and how the artistic unity that we vaguely feel is actually created. Generally, artworks are often more interesting and perceived as ‘deeper’ if they allow for many different interpretations, because this will make us think longer about them, and invest more creative energy into the interpretation process. This is also true for intermodal relations in artworks, which wouldn’t be interesting if they were too obvious.

## V.4 From individual to social style

In his book *Introducing Social Semiotics*, Theo van Leeuwen presents an approach to semiotic analysis which distinguishes four dimensions, namely discourse, genre, style, and modality (van Leeuwen 2005: 91). The four concepts complement each other and can be understood to commonly define a situated approach to multimodal communication. While *discourse* concerns the use of semiotic resources to construct representations of aspects of the world (i.e. what in structuralist terms would be called *parole*, the use of signs for specific communicative purposes), *genre* refers to the type and specific conventions of the multimodal communication process. On this basis, style is understood as follows:

The concept of ‘style’ is the key to studying how people use semiotic resources to ‘perform’ genres, and to express their identities and values in doing so. (van Leeuwen 2005: 91)

Van Leeuwen (ibid.: 140) distinguishes three different concepts of style, namely individual style, social style, and lifestyle. ‘Individual style’ concerns the ways in which a person expresses her- or himself multimodally through their specific ways of using semiotic resources:

The individual style of a voice allows you to recognise who the speaker is, the individual style of handwriting who the writer is, and so on. [...] For others, style does have meaning – expressive meaning. Style expresses feelings, attitudes to what is said – or written or painted or acted – and it expresses the personality of the speaker – or writer, or painter or actor. (van Leeuwen 2005: 140)

The passage makes it clear that styles are signs, in the sense that they allow perceivers to draw inferences about who created a text or artefact, and about their feelings and attitudes. Apart from inferences about personality and education of the individual style producer or user, we can also draw inferences about their social background:

The idea of ‘social style’ foregrounds [...] that style expresses, not our individual personality and attitudes, but our social position, ‘who we are’ in terms of stable categories such

as class, gender and age, social relations, and 'what we do' in terms of the socially regulated activities we engage in and the roles we play within them. (van Leeuwen 2005: 143)

'Social style' also refers to sign functions of styles, but in this case related to social aspects of the person in question. The categories of individual style and social style can therefore be regarded as complementary. While the notion of individual style foregrounds the personality and individual motivation of communication and behaviour, the notion of social style focuses on the social factors that contribute to it. In this perspective, stylistic characteristics are regarded as markers for the variety found in modern societies (Siefkes 2013).

Both aspects of styles play an important role in everyday life. In a recent study (Siefkes 2017; cf. section II.4.3), one of the authors of this book investigated the connection between style perception and the ability to attribute novel excerpts. It turned out that acute perception of stylistic features weakly predicted the subjects' ability to identify those novel excerpts that belonged to the same novel. Effects such as these might seem obvious, but there has been little experimental confirmation for them. Certainly, there can be little doubt that experts (for example art historians or archaeologists) are able to gather knowledge from the style of a painting or an artefact – for example to ascertain their place and time of origin, social functions, and influences from other cultures, and in some cases to guess who the artist is. However, these abilities may be limited to experts who are well-versed in a specific field and have extensive experience with the artefacts in question. Further research should ascertain to what degree laypersons are able to gather knowledge from styles.

On the other hand, styles of speech and social interaction seem to be widely understood. In daily interactions, most of us are wont to make guesses about the personality and social background of our interlocutors, based on their styles of speech, gesture, body posture, proxemics, and gaze behaviour, as well as on their general style of social interaction (for example how they introduce themselves). Obviously, we are confronted with such styles every day, but it is unclear whether this provides us with the extensive knowledge and deep understanding about them that is characteristic for expertise. The necessity of expertise for the perception and interpretation of various types of style will require more detailed studies.

The third style category introduced by van Leeuwen (2005: 144–148) is 'lifestyle', which combines individual and social style. The concept of lifestyle can be justified by the fact that styles of dressing, consumer choices, leisure-time activities, and attitudes towards key social issues are often correlated with each other to some degree, which makes it plausible to assume an integrative level that may be called a 'lifestyle' (cf. Chaney 1996).

In fact, the many styles an individual uses in different areas of artefact use and behaviour are to some degree indicative of each other, confirming the general notion of 'lifestyle'. For example, eating choices hint towards other preferences, such as favourite sports or cultural interests. These connections are of course abundantly exploited by the advertisement and marketing sector. For example, in the same cinema, the audience of a blockbuster movie will be targeted with different commercials than the viewers of an independent movie. Such differentiation is successfully employed in various areas such as food, sports, travel, and entertainment. The success of selective targeting has been empirically validated in marketing studies.

It should be noted that the widespread and successful targeting of advertisements on the basis of lifestyle choices not directly related to the product advertised, confirms the long-standing semiotic approach to styles which assumes that styles contain meanings and can be interpreted (Chatman 1971; Thoma 1976; Trabandt 1979; Bayer 1989; Spillner 1995; Hess-Lüttich/Wenz 2006; Siefkes 2011, 2012a). If lifestyle choices and preferences in different areas are to some degree connected, it is obvious that knowledge about an individual's style in one specific area enables educated guesses about preferences and choices in other areas.

In a more general perspective, style can be understood as a category of analysis that links social conventions with individual behaviour. In fact, style is traditionally applied in three different ways: it can be used to characterize specific texts or artefacts (*text/artefact style*), producers (*individual style*), or groups of producers (*group style*). In other words, it is possible to either analyse the stylistic qualities of an individual text or artefact, or to analyse the style of a specific author, musician, or artist.

For example, we can analyse the style of *2001: A Space Odyssey* (1968) by just focusing on this film, or in relation to Stanley Kubrick's other films. In fact, Kubrick's films show an impressive stylistic diversity and we may question whether an 'individual style' can be assigned to Kubrick. It may even be argued that he adapted and combined existing styles or developed new stylistic qualities for the various films he made. For other film-makers, like Hitchcock, the assumption of an individual style (obviously with changes over time) is more plausible.

However, a style can also be assigned to groups of writers or artists, to social groups such as a subculture, a class (e.g. 'proletarian', 'aristocratic', or 'elitist' style), a profession (with its specific lingo, clothing, and lifestyle), or a conventionalized situation (such as a sermon or a trial). Therefore, style is a category that is used to describe both individual characteristics and social differences, and may be understood as a bridge between individual and social behaviour.

## V.5 Style in the media

In their book *Global Media Discourse*, David Machin and Theo van Leeuwen analyse the style of international media. In one chapter, they compare the British edition of *Cosmopolitan* to the various international editions of this successful women's magazine, and analyse the way the magazine creates identity by telling stories (Machin/van Leeuwen 2007: 41–73; cf. also van Leeuwen 2005: 148–159, where the US, Dutch, Spanish, Indian, and Chinese versions of *Cosmopolitan* are compared).

*Cosmopolitan* uses a number of more or less conventionalized styles:

The principal styles on which *Cosmopolitan* draws are (1) the style of advertising, (2) the style of the fashion caption, (3) the style of expert discourse, (4) street style – the slang of the trendy, and the young, and (5) conversational style. These mix to different degrees with traditional magazine feature writing style – the traditional, socially 'appropriate' style for magazines – in the same way that lifestyle dress may be combined with, or tempered by, for instance, traditional 'white collar' work dress. (van Leeuwen 2005: 149)

These styles are combined in different ways depending on the specific topic and functions of the article in question. Furthermore, the stylistic combinations and hybrid forms differ between the various local versions of the magazine. Interestingly, the analysis of *Cosmopolitan* and other internationally franchised media shows that linguistic style is more localized, allowing for local 'accents' of the globalized product. For visual style, on the other hand, economic and technical factors make localization more difficult (Machin/van Leeuwen 2007: 170–171). This difference leads to more or less pronounced intermodal style discrepancies. Whereas the original US version of *Cosmopolitan* could probably be described as a stylistic unity, the local versions resemble the US version more in their visual style than in the style of their language.

One example can be seen in the Spanish version of *Cosmopolitan*, which mostly avoids informal style and 'street language' which are still seen as uneducated in Spain. In comparison, the Chinese version emulates the original's informal style:

The Chinese version, on the other hand, has adopted the Cosmo style with enthusiasm: 'Let's compete to see who is more joyfully casual.' There are economic reasons for this, because they see it as a style that will attract advertisers. At the same time, in creating a Chinese version of Cosmo 'poetics' they draw on classical Chinese styles, such as the symmetrical arrangements of words in the 'antithetical couplet', rather than on Western poetic devices. Apparently market reform has been accompanied by a revitalisation of traditional forms in China. (Machin/van Leeuwen 2007: 143)

The Chinese version's writing style therefore resembles the US version more closely than the Spanish version's style. This can be related to the positive cultural evaluation of slang and informal language in countries such as the US and China, where they are associated with coolness or hipness and generally with a modern lifestyle.

On the other hand, the images and layout of the franchises generally resemble the US version's glossy, glamorous visual style, according to Machin's and van Leeuwen's analysis. It would be worthwhile to study possible tensions between writing style and visual style that might arise from this 'glocalization' (localization of a global product; cf. Sigismondi 2011) of international media franchises. In any case, it is possible that readers adapt to the specific mixture of linguistic and visual style, perceiving it as characteristic for the overall style of a magazine or other medium, and would only notice discrepancies if they were to compare different editions.

The example shows that it can be worthwhile to compare styles in different modes. It is often taken for granted that the style, for example, of magazines, websites, or films can be understood in an integral, holistic perspective. While it is plausible that style has some holistic effects, which cannot be reduced to combinations of stylistic effects of the participating modes, such examples show that the analysis of style can profit from a perspective that explicitly considers the contributions of the various modes towards a 'multimodal style'. For example, a print journal or website can be investigated in regard to the style of its images, verbal text, layout, and typography. A face-to-face interaction can be analysed in regard to the style of speech, of gesture, of facial expression, body posture, and proxemics. For a film, the style of moving images, speech, and music can be distinguished, as well as further modes embedded in the images (such as gesture, posture, and facial expression of the characters). We can then look for stylistic qualities of the separate modes, or for congruence or incongruence of the style of the various modes.

We might also look for higher-level effects: e.g. stylistic references to different epochs, or intertextual citations, which appear in only one mode, but modify the meaning of the others. To take a fictitious example: think of an action film's music that cites Ennio Morricone's famous soundtracks, thus giving the film a touch of the 'spaghetti western', which might alter the way the images or character's speech are perceived.

In multimodality research, it has been stressed that different modes vary in regard to their affordances (Kress 2014: 61), which is usually spelled out to imply that they are differently suited for expressing specific meanings. However, the stylistic affordances of modes can also vary. For example, styles of images enable the foregrounding of a certain understanding of what is shown, in ways that are

difficult for language. An interesting example is given in van Leeuwen (2008: 145), where two different images of Muslim women are shown that originally appeared in an article in the *The Guardian Weekend* in 2001.

One of the images shows a group of women seen from above, they are 'looked down upon' (ibid.: 146), nearly all of them wear hijabs or veils, a fact which is foregrounded by the camera perspective, and all of them look very serious and raise their eyes towards the observer as if in supplication. The other image also shows a number of Muslim women, but with a much larger variety of attires, facial expressions, and skin colours; all of them are portrayed frontally and look towards the viewer from the same height level. Since both images have basically the same content (they show a number of Muslim women), the differences could be characterized as primarily stylistic, consisting in the angle of view, and the way the women are presented. In one case, Muslim women are presented as basically similar, in the other case, as individuals with different lives, personalities, and things to communicate.

In short, this article juxtaposes two different ways of looking at Muslim women – one in which they are represented as equals and brought close to 'us', Guardian readers, and one in which they are homogenized and looked down upon. (van Leeuwen 2008: 146)

The text of the article explicitly considers two different perspectives on Muslim women, as either simply suppressed by their religion, or as agents that have a degree of freedom in their actions, and have to be considered with their individual characteristics and positions in society. The contrasting visual styles therefore serve to underline the two perspectives explicitly spelled out in the text of the article – an interesting and relevant intermodal use of style.

This result can be generalized. In many cases, the individual style of a journal, magazine, TV show, or other multimodal media type will include specific uses of the separate modes. Images can be used to supply additional information, to provide atmosphere and arouse interest, or as a visual embellishment of the text. Graphics may be employed to illustrate what the text already explains, or in a more independent manner, supplying information and additional detail. Typography can be used mainly to structure the text, or to supply a certain feeling and underline specific values (such as clarity or elegance) of the publication in question. The list could be continued *ad infinitum*.

The uses a multimodal medium makes of the various modes it employs – how it distributes the content across modes and uses them to fulfil various communicative functions – distinguishes it from other media of the same type (e.g. other journals, magazines, or TV shows). Therefore, these aspects should be regarded as part of the style of a multimodal text.

## V.6 Multimodal style and genre

It is plausible to assume that the stylistic qualities of a multimodal text depend to some degree on the *genre* of the text in question (cf. Bateman 2008). Generally, style and genre are related, but not identical descriptive categories, since the genre of a text determines to some degree the stylistic qualities that we expect from it (Siefkes 2012a: 412). Think of a newspaper article, a secondary-school textbook, an article on a private blog, and a legal text: while these four genres may all be used to talk about a specific topic (such as racism, alcoholism, or police brutality), the stylistic qualities we expect of the respective texts vary substantially.

Some of these stylistic expectations are motivated by the specific textual functions that the text genre fulfils, as well as the audience it is directed at, the institutional background of its production and use, and the typical situations in which it occurs. For example, a legal text has the function to clarify the legal side of the respective problem, will be primarily directed at law professionals, and will be used in the context of courts, law schools, and other legal institutions in communicative situations such as lawsuits or teaching situations. However, there are other stylistic conventions which are not necessary to fulfil the functions of the genre in question (though we don't always realize this). For example, we expect a legal text to be dry and complicated 'legalesé', while these properties are by no means necessary (and may actually be detrimental) for these texts to fulfil their genre-related functions. These latter expectations are caused by 'conventions' (Lewis 1969), principles that are not determined by necessity, but which have nonetheless emerged as 'the usual way of doing things'. Obviously, formulating legal problems in a clear and easily readable manner has not yet become a general convention for legal texts – although there are certainly specific subgenres, such as certain websites explaining legal problems, which aim for a clear and understandable language.

The same can be said for many multimodal text genres, but here the presence of various modes with their specific stylistic qualities adds an additional level of complexity. For example, we may expect certain qualities from the language of a newspaper article or a textbook, but we may also have expectations for the use of images or graphics, or for the typography suitable for these genres. It is plausible that such conventions even include the 'distribution of content over modes', specifying which parts of the message are expressed in which mode. For example, we would expect a newspaper article of a crime to show a picture of the crime scene or maybe of the victim, but would probably be surprised to see a panoramic view of the town where the crime took place, or a photo of the

prosecutor. We would rather expect the town and the name of the prosecutor to be mentioned in the text. Behind such expectations, there are conventions about which mode is aptly used for which aspect of the story – while it is seen as legitimate that the public may be interested in visuals of the crime scene,<sup>17</sup> it isn't supposed to know how the prosecutor looks, presumably out of respect for his or her privacy. However, the public supposedly has the right to know the name of the prosecutor, who serves in a public position and has an official role in the trial. Similar conventions and restrictions can be found for many multimodal genres.

In order to understand the role of the different modes in a multimodal genre, one first has to analyse the genre-related functions of a specific multimodal text genre. Genre-related functions are 'what the text has to do', for example to tell a story, create reader engagement, impart emotions, provide background information, be understandable to a certain target group, etc. Generally speaking, the stylistic qualities of multimodal texts depend to some degree on the functions related to the respective genre, and on conventions which determine 'what is done by which mode' in order to fulfil these functions.

Sandgren (2010) provided an example of this in action, where the relations between text, images, and maps are investigated for a corpus of bird guides. Sandgren points out that the semiotic modes used in bird guides – text, images, maps, etc. – typically show certain stylistic qualities, which can only in part be explained with the functions of the genre (namely, to identify birds):

The style in field guides is matter-of-fact and the text has to act together with the images to make identification possible. This does not stop the author from using rhetorical tricks, sometimes quite artfully, or using narratives. [...] The text is often supported by a map which shows the ranges of the species, during breeding season or during migration. To serve its goal, the text is primarily neutral and descriptive. (Sandgren 2010: 114)

In highly specialized genres, the specific functions of the text are often to some degree divided between the participating semiotic modes. In the case of bird guides, the images are central for fulfilling the primary textual function of bird identification. Detailed and high-quality images are used, which, on the other hand, show prototypical specimens in generic and low-key contexts. A reader not acquainted

---

17 An example from the German media: in a recent case of kidnapping and murder that received extensive news coverage, many articles on news websites showed the house where the crime took place (e.g. <http://www.spiegel.de/panorama/justiz/hoexter-dasist-ueber-die-verdaechtigen-wilfried-w-und-angelika-w-bekannt-a-1090576.html> [accessed 12 December 2016]). Even though the pictures just show an average, slightly dilapidated house, it is obviously assumed that a legitimate curiosity is satisfied by showing them.

with the genre might perceive a contradiction between the glossy presentation and the somewhat generic content of the bird images – in comparison, let's say, to images of birds in a coffee-table book. The contradiction vanishes if a reader is acquainted with the genre and its specific textual functions. These functions imply constraints both for the content and for the mode of presentation.

Similarly, the mentioned stylistic qualities of the text ('matter-of-fact', 'neutral', 'descriptive') are, at least to some degree, prompted by functional demands. The example shows that multimodal genres often have specific constellations both of content and function of the participating semiotic modes (such as images, text, maps, and layout/typography), which also influence the stylistic aspects of these modes.

We can therefore expect to find specific styles for all participating modes, as well as specific intermodal stylistic relations, for multimodal text genres such as field guides (Sandgren 2010), tourist brochures (Hiippala 2015), or comics (Cohn 2015). Generally speaking, multimodal genre analysis provides a range of methods for empirical analysis of multimodal text genres. For page-based documents, the genre and multimodality model (Bateman 2008; Hiippala 2014) provides an approach using XML-based tag sets, which enables researchers to annotate multimodal corpora in such a way that automatic search queries become possible. This allows for a precise annotation and statistical evaluation of multimodal corpora.

Another approach towards quantitative, corpus-based research in multimodality was developed by Doris Schöps (2013, 2016). A corpus of 75 DEFA (Deutsche Film AG) feature films was annotated with the software ELAN (Wittenburg et al. 2006), and evaluated with statistical methods. The study is based on the assumption that body postures are a sign system (Schöps 2016: 28) that possesses communicative functions. In the terminology used in this chapter, sign systems can roughly be equated to semiotic modes. While body postures in film are often not explicitly considered as a separate semiotic mode, they can be characterized as a mode that is embedded in the more general mode 'moving images'. On the basis of a detailed description of 42 different body postures (Schöps 2016: 155–257), the study investigates the use of body postures in DEFA feature films in relation to the characters that adopt them and to the genre of the film (e.g. adventure & entertainment vs spy & crime thriller).

It is shown that body postures have a number of functions in films, and that the use of body postures is connected to characteristics both of the genre and of the narrative structure of the respective films. While this study specifically investigates the functions of body postures in DEFA films, the developed method,

which combines principles of corpus construction, annotation, and statistical evaluation, can be applied to other multimodal text genres as well. Furthermore, it enables new and more precise approaches to a whole range of research questions in film studies, as well as other areas of multimodality research (Schöps 2016: 456–457). An English summary of this study can be found in Schöps/Siefkes (in preparation).